

WHAT IS CLAIMED IS:

1 1. A system for determining popularity of web pages on a network, the
2 system comprising:
3 a plurality of monitoring devices placed in the network, the monitoring
4 devices monitoring packets traversing the network and extracting
5 information on the packets; and
6 a processing module coupled to the monitoring devices and receiving the
7 extracted information from the monitoring devices, the processing
8 module analyzing the extracted information and determining the
9 popularity of the web pages based upon the extracted information,
10 the popularity of the web pages being substantially proportionate
11 to number of visits to the web pages as indicated by the extracted
12 information.

1 2. The system of claim 1, wherein the monitoring devices are placed in
2 locations where aggregate packet traffic may be monitored.

1 3. The system of claim 1, wherein the monitoring devices are placed at a
2 traversal point for complete activity between a client device and a server on the network.

1 4. The system of claim 1, wherein the monitoring devices extract the
2 information from packets in a TCP session, and the extracted information includes:
3 a requested URI or URL;
4 a client IP address; and

5 a server IP address and a server host name.

1 5. The system of claim 4, wherein the extracted information further includes
2 a referrer URL.

1 6. The system of claim 4, wherein the monitoring devices analyzes the
2 packets relating to GET Requests in the TCP session to extract the information.

1 7. The system of claim 1, wherein the monitoring devices discard packets
2 relating to invalid URLs, invalid GET Requests, requests from a web crawler, or auto-
3 refreshment of previous TCP sessions in extracting the information.

1 8. The system of claim 1, wherein the processing module maintains a counter
2 corresponding to a URL and increments a count of the counter, if the extracted
3 information indicates that the web page corresponding to the URL was visited, the count
4 indicating the number of visits to the web page.

1 9. The system of claim 1, wherein the processing module maintains a
2 plurality of counters corresponding to a URL and increments a count of one of the
3 counters, if the extracted information indicates that the web page corresponding to the
4 URL was visited by a client device located in a geographical location corresponding to
5 the counter of which the count was incremented, the count indicating the number of visits
6 to the web page by client devices in the corresponding geographical location.

1 10. The system of claim 9, wherein the processing module increments the
2 count only if the extracted information indicates that the web page was visited by the
3 client device having a distinct IP address.

1 11. The system of claim 9, wherein the processing module does not increment
2 the count if the extracted information indicates that the packets were automatically and
3 repeatedly generated by a computer.

1 12. A search system for ranking Internet search results based upon popularity
2 of web pages on a network, the search system comprising:

3 a plurality of monitoring devices placed in the network, the monitoring
4 devices monitoring packets traversing the network and extracting
5 information on the packets;

6 a processing module coupled to the monitoring devices and receiving the
7 extracted information from the monitoring devices, the processing
8 module analyzing the extracted information and determining the
9 popularity of the web pages based upon the extracted information,
10 the popularity of the web pages being substantially proportionate
11 to number of visits to the web pages as indicated by the extracted
12 information; and

13 a search engine for receiving search terms and retrieving web pages
14 containing the search terms, the search engine ranking the web
15 pages at least in part based upon the popularity of the retrieved
16 web pages.

1 13. The search system of claim 12, wherein the search engine ranks the
2 retrieved web pages based upon the content of the web pages and the hyperlink structure
3 linking the web pages as well as the popularity of the retrieved web pages.

1 14. The search system of claim 12, wherein the search engine propagates a
2 score of a first web page to a plurality of second web pages to which the first web page is
3 linked substantially in relative proportion to the popularity of links from the first web
4 page to each of the second web pages.

1 15. The search system of claim 12, wherein the search engine ranks a first
2 retrieved web page in higher priority than a second retrieved web page if the popularity
3 of the first web page is greater than the popularity of the second web page.

1 16. The search system of claim 12, wherein the monitoring devices are placed
2 in locations where aggregate packet traffic may be monitored.

1 17. The search system of claim 12, wherein the monitoring devices are placed
2 at a traversal point for complete bi-directional activity between a client device and a
3 server on the network.

1 18. The search system of claim 12, wherein the monitoring devices extract the
2 information from packets in a TCP session, and the extracted information includes:

3 a requested URI or URL;

4 a client IP address; and

5 a server IP address and a server host name.

1 19. The search system of claim 18, wherein the extracted information further
2 includes a referrer URL.

1 20. The search system of claim 18, wherein the monitoring devices analyzes
2 the packets relating to GET Requests in the TCP session to extract the information.

1 21. The search system of claim 12, wherein the monitoring devices discard
2 packets relating to invalid URLs, invalid GET Requests, requests from a web crawler, or
3 auto-refreshment of previous TCP sessions in extracting the information.

1 22. The search system of claim 12, wherein the processing module maintains a
2 counter corresponding to a URL and increments a count of the counter if the extracted
3 information indicates that the web page corresponding to the URL was visited, the count
4 indicating the number of visits to the web page.

1 23. The search system of claim 12, wherein the processing module maintains a
2 plurality of counters corresponding to a URL and increments a count of one of the
3 counters if the extracted information indicates that the web page corresponding to the
4 URL was visited by a client device located in a geographical location corresponding to
5 the counter of which the count was incremented, the count indicating the number of visits
6 to the web page from client devices in the corresponding geographical location.

1 24. The search system of claim 23, wherein the processing module increments
2 the count only if the extracted information indicates that the web page was visited by the
3 client device having a distinct IP address.

1 25. The search system of claim 23, wherein the processing module does not
2 increment the count if the extracted information indicates that the packets were
3 automatically and repeatedly generated by a computer.

1 26. The search system of claim 12, wherein the monitoring devices detect
2 requests to stale web pages.

1 27. The search system of claim 12, wherein the monitoring devices detect
2 pages unknown to the search engine.

1 28. A method for determining popularity of web pages on a network, the
2 method comprising:
3 receiving a TCP packet traversing the network;
4 determining from the received TCP packet whether a TCP session has
5 started; and
6 responsive to determining that a TCP session has started, extracting
7 information on the TCP session from subsequent TCP packets
8 traversing the network, the extracted information indicating the
9 popularity of the web pages, and the popularity of the web pages
10 being substantially proportionate to number of visits to the web
11 pages as indicated by the extracted information.

1 29. The method of claim 28, wherein receiving a TCP packet comprises
2 retrieving packets having a protocol field value of 6.

1 30. The method of claim 28, wherein determining whether a TCP session has
2 started comprises determining whether the received TCP packet is a SYN packet and has
3 a destination port number of 80.

1 31. The method of claim 28, wherein the extracted information includes:
2 a requested URI or URL;
3 a client IP address; and
4 a server IP address and a server host name.

1 32. The method of claim 31, wherein the extracted information further
2 includes a referrer URL.

1 33. The method of claim 31, further comprising increasing a popularity count
2 corresponding to a web page if the extracted information indicates that the web page was
3 visited as determined based upon the requested URI or URL.

1 34. The method of claim 33, wherein the popularity count indicates the
2 number of visits to the web page from client devices in a geographical location
3 corresponding to the popularity count.

1 35. The method of claim 33, wherein the popularity count is increased only if
2 the extracted information indicates that the web page was visited by a client device
3 having a distinct IP address.

1 36. The method of claim 33, wherein the popularity count is not increased if
2 the extracted information indicates that the packets were automatically and repeatedly
3 generated by a computer.

1 37. The method of claim 28, wherein extracting information on the TCP
2 sessions comprises discarding packets relating to invalid URLs, invalid GET Requests,
3 requests from a web crawler, or auto-refreshment of previous TCP sessions.

1 38. A method for ranking Internet search results based upon popularity of web
2 pages, the method comprising:

3 receiving a search term;

4 performing search of web pages on the Internet based upon the received

5 search term;

6 retrieving a plurality of web pages containing the search term; and

7 ranking the web pages at least in part based upon the popularity of the

8 retrieved web pages, the popularity of the retrieved web pages

9 being substantially proportionate to number of visits to the web

10 pages.

1 39. The method of claim 38, wherein ranking the web pages comprises
2 ranking the web pages based upon the content of the web pages and the hyperlink
3 structure linking the web pages as well as the popularity of the retrieved web pages.

1 40. The method of claim 38, further comprising propagating a score of a first
2 web page to a plurality of second web pages to which the first web page is linked

3 substantially in relative proportion to the popularity of links from the first web page to
4 each of the second web pages.

1 41. The method of claim 38, wherein ranking the web pages comprises
2 ranking a first retrieved web page in higher priority than a second retrieved web page if
3 the popularity of the first retrieved web page is greater than the popularity of the second
4 retrieved web page.

1 42. A system for determining popularity of web pages, the system comprising:
2 a plurality of monitoring means placed in a network for monitoring
3 packets traversing the network and extracting information on the
4 packets; and
5 a processing means coupled to the monitoring means for receiving the
6 extracted information from the monitoring device, analyzing the
7 extracted information, and determining the popularity of the web
8 pages based upon the extracted information, the popularity of the
9 web pages being substantially proportionate to number of visits to
10 the web pages.

1 43. A search system for ranking Internet search results based upon popularity
2 of web pages, the search system comprising:
3 a plurality of monitoring means placed in a network for monitoring
4 packets traversing the network and extracting information on the
5 packets;

6 processing means coupled to the monitoring means for receiving the
7 extracted information from the monitoring devices, analyzing the
8 extracted information, and determining the popularity of the web
9 pages based upon the extracted information, the popularity of the
10 web pages being substantially proportionate to number of visits to
11 the web pages; and
12 search engine means for receiving search terms and retrieving web pages
13 containing the search terms, the search engine means ranking the
14 web pages at least in part based upon the popularity of the
15 retrieved web pages.

1 44. A method for determining popularity of links from a first web page to a
2 plurality of second web pages on a network, the method comprising:
3 receiving TCP packets traversing the links;
4 extracting information on a TCP session comprised of the received TCP
5 packets; and
6 determining popularity of the links based upon the extracted information,
7 the popularity of each of the links being substantially proportionate
8 to number of times each of the links is traversed.

1 45. The method of claim 44, further comprising propagating a score of a first
2 web page to a plurality of second web pages to which the first web page is linked
3 substantially in relative proportion to the popularity of links from the first web page to
4 each of the second web pages.

1 46. The method of claim 44, wherein the extracted information includes at
2 least a referrer URI and a requested URI of the TCP packets.

1 47. A system for determining popularity of links from a first web page to a
2 plurality of second web pages on a network, the system comprising:

3 a plurality of monitoring devices placed in the network, the monitoring
4 devices monitoring packets traversing the network and extracting
5 information on the packets; and

6 a processing module coupled to the monitoring devices and receiving the
7 extracted information from the monitoring devices, the processing
8 module analyzing the extracted information and determining the
9 popularity of the links based upon the extracted information, the
10 popularity of each of the links being substantially proportionate to
11 number of times each of the links is traversed.

1 48. A search system comprising:

2 a plurality of monitoring devices placed in the network, the monitoring
3 devices monitoring packets traversing the network and extracting
4 information on the packets;

5 a processing module coupled to the monitoring devices and receiving the
6 extracted information from the monitoring devices, the processing
7 module analyzing the extracted information and determining the
8 popularity of the links from a first web page to a plurality of
9 second web pages based upon the extracted information, the

10 popularity of each of the links being substantially proportionate to
11 number of times each of the links is traversed; and
12 a search engine for receiving search terms and retrieving web pages
13 containing the search terms, the search engine propagating a score
14 of the first web page to the second web pages to which the first
15 web page is linked substantially in relative proportion to the
16 popularity of links from the first web page to each of the second
17 web pages.